The Southeast US as a regional model for global studies

Zack Murrell
Appalachian State University

Using Biodiversity Specimen-Based Data to Study Global Change Workshop
Missouri Botanical Garden, St. Louis
2-3 December 2015
Why: A biodiverse region with intense human population pressures

Hot spots of plant endemism in the southeast (Estill & Cruzan 2001)
Electronic communication and storage now allows us to move information at breathtaking speeds across space and time, opening up new opportunities to aggregate and parse large parcels of information.
"As access to information dramatically expands, so that people increasingly have access to almost all the information they might need at any time and in any place (and, surprisingly, at low or no cost), the value of the cognitive skills still unreplicable by silicon becomes greater."

But this value cannot be recognized without providing appropriate access for biologists.

Cyberinfrastructure

A combination of databases, network protocols and computational services that brings people, information and computational tools together to perform science in this information-driven world.

Lincoln D. Stein

www.nature.com/reviews/genetics
Cyberinfrastructure

• Data infrastructure and storage
• Computational infrastructure
• Communication infrastructure
  Low-level (connectivity and bandwidth)
  Syntactic (common formats to organize data)
  Semantic (common terminology/ontology)
• Human infrastructure
  Sharing and curation

Lincoln D. Stein

www.nature.com/reviews/genetics
We are drowning in information, while starving for wisdom. The world henceforth will be run by *synthesizers*, people able to put together the right information at the right time, think critically about it, and make important choices wisely.

E.O. Wilson

*Consilience: The Unity of Knowledge* (1998)
Selfish and contentious people will not cohere, and without coherence nothing can be effected.

Charles Darwin  Descent of Man.  1871
Principles of the Conservation Commons

1. Promote Free and Open Access

2. Mutual Benefit (Reciprocity)

3. Rights and Responsibilities
   - Attribution
   - Integrity of Original Work

http://www.conservationcommons.net/
The Key to the Cabinets: Building and Sustaining a Research Database for a Global Biodiversity Hotspot

A new NSF-Advancing Digitization in Biological Collections (ADBC) collaborative project in the southeast

NSF ADBC #1410069 (plus collaborative awards 1410077 1410081 1410086 1410087 1410092 1410094 1410098 1410143 1410200 1410288 1410439 1410445)
• 2005 NSF Research Coordination Network (RCN) awarded to build the ‘human infrastructure’ to begin the ‘2020 Vision effort’
• 5-year project
• 150 herbaria participated

Used existing infrastructure of the Association of Southeastern Biologists
What is a herbarium?

Organized collections of carefully preserved and documented plant specimens

-plus-

Associated resources such as curators, taxonomic specialists, collections managers, students, floras, monographs, microscopes, and curation supplies.
State of Herbaria in the Southeast USA

- 233 collections vary in size from 3,000 to 750,000 specimens.
- Curators are taxonomists, ecologists, paleobotanists, population biologists.
- Some institutions with significant budgets, IT support and curatorial staff.
- Some institutions are single person operations without budgets or IT support.

- ESTIMATED 15 MILLION SPECIMENS FROM THE SOUTHEAST.
Who is involved in the ADBC project?

- 93 herbaria (plus 12)
- IT (Symbiota, GEO Locate, SPECIFY, iPLANT-TACC)
- Citizen scientists
  - Notes From Nature
  - Adler Museum
- Education experts
- Herbarium Affiliates
- State Natural Heritage Programs
- iDigBio
Data Management Workflow

Specimen Image Capture “Herbarium Digitarium”

4,700,000 SPECIMENS
100+ INSTITUTIONS
12 STATES

Texas Advanced Computing Center

Archival Storage

Web Derivatives

Citizen Scientists

Metadata to Portal

Batch Georeferenced Data to Portal

Repatriation of Metadata to Local Institution
Symbiota and Specify

GEOLocate

Transcribe museum records to
TAKE NOTES FROM NATURE
START TRANSCRIPTION

SERNEC
Southeast Regional Network of Expertise and Collections

Home >> Collections

Collections to be Searched

Specifications & Observations
Specimens
Observations
Federal Units

Appalachian State University (BSU)
Virginia State University Herbarium (VSUH)
A. C. Moore Herbarium - University of South Carolina (USC)
Arkansas Tech University Herbarium (APCH)
Auburn University - John D. Freeman Herbarium (AUA)
Clemson University Herbarium (CLES)

Take notes from nature.
MOBILIZATION AND COMMUNICATION

WORKING GROUPS:
1. IT and infrastructure
2. Libraries and literature
3. Images (specimen and live)
4. Education and outreach
5. Taxonomic concept mapping
6. State working groups

“IT TAKES A VILLAGE”
Steps to increasing the value of natural history collections

Collection → GPS → ID → Label → Mounted specimen

- Digitize and image
- Mapping
- Modelling
- Phenology
- Micro-morphology
- Herbivory
- DNA
- Concept maps

Quality Assurance/Quality Control at all steps
Governance structure based on the Society of Herbarium Curators constitution

State consortia

Regional chapters of SHC

Decision-making
Consensus building
Conflict resolution
Knowledge transfer
Longevity
Sustainability

http://www.societyofherbariumcurators.org
What are the ADBC project goals?

Use our robust human infrastructure

1) improve protocols at every stage of specimen digitization.

2) interface with citizen scientists to develop a strong collaborative community that will sustain high transcription and georeferencing rates through a multi-year effort.
What works?

1. Use existing infrastructures (Association of Southeastern Biologists)
2. Provide a reward system of making the group’s aggregated data available for their own manipulation
3. Transformational leadership?
4. Communication???
What can biologists do?

1. Build our communities (include amateurs and affiliates).
2. Embrace and promulgate new technologies.
3. Take part in the development of community best practices (among data creators, users, editors and publishers).
4. Expand the community to include computer, social and library scientists.
5. Generate a paradigm shift in how productivity is evaluated by administrators.
6. Utilize more effective means of communication.

Snapshot versus long-term shift in best practices. We need to train biologists in traditional studies of natural history, while at the same time encouraging the use of new technologies.
SERNEC: SouthEast Regional Network of Expertise and Collections

National Science Foundation
Research Coordination Network
RCN Award # 0542320

National Science Foundation
Advancing Digitization of Biological Collections
Award # 1410069